

# A Data Quality aware Enterprise Service Bus for e-Health Integration Platforms

Silvana Pidre, Laura González, Rodrigo Mendoza, Mauricio Piñatares, Nicolas Granja, Flavia Serra and Raúl Ruggia

Instituto de Computación, Facultad de Ingeniería

Universidad de la República

Julio Herrera y Reissig 565, Montevideo, Uruguay

Email: {spidre, lauragon, fserra, ruggia}@fing.edu.uy

**Resumen**—The integration of software systems running in different healthcare organizations is increasingly required in order exchange data and carrying out healthcare processes in coordinated ways. This integration is usually supported by integration platforms which are based on advanced middleware technologies, like the Enterprise Service Bus (ESB). In addition, data quality in the healthcare area is critical to take timely and accurate decisions for the health and well-being of patients. This paper proposes a data quality aware ESB (DQ-ESB) for e-Health integration platforms which allows assessing and enforcing data quality within inter-organizational data exchanges. DQ-ESB handles an extensible data quality model which includes quality characteristics identified within the Uruguayan e-Health context. DQ-ESB provides data quality solutions based on native ESB mechanisms (e.g. data transformation) and it was implemented leveraging the Switchyard ESB product.

**Index Terms**—Data Quality, Enterprise Service Bus, e-Health, Integration Platform.

## I. INTRODUCCIÓN

La integración de sistemas de software operando en diferentes organizaciones de salud es cada vez más necesaria para intercambiar datos y llevar a cabo procesos de atención médica de manera coordinada [1]. Esta integración es fundamental, por ejemplo, para componer la historia clínica completa de un paciente que puede estar distribuida en diferentes organizaciones de salud (p. ej. laboratorios, prestadores, ministerio).

Una forma de abordar la integración de sistemas de *software* de diferentes organizaciones es a través de plataformas de integración [2]. Estas plataformas son infraestructuras especializadas que brindan una capa de procesamiento intermedio entre las organizaciones para facilitar problemáticas de integración (p. ej. conectividad, heterogeneidad de modelos y formatos de datos) [2] y se basan en tecnologías de *middleware* avanzadas, como el Bus de Servicios Empresariales (*Enterprise Service Bus*, ESB) [3]. En áreas como el gobierno [4] y la salud [5], estas plataformas pueden ser provistas por entidades gubernamentales como parte de marcos más amplios de interoperabilidad. Estos apuntan a facilitar y promover la interconexión entre las organizaciones brindando servicios que generan economía de escala [4].

Por otro lado, la calidad de datos en el área de la salud es fundamental para tomar decisiones oportunas y precisas para el bienestar y la salud de los pacientes. La importancia de

este tema se debe al impacto que puede tener en la mejora de la atención a los pacientes, en la definición de políticas de salud para la población y en la gestión de los fondos destinados para el mantenimiento de los servicios de salud [6]. Para poder mejorar la calidad de los datos recopilados, y de la información generada a partir de estos, es necesario adoptar medidas de control de la calidad [6]. Uno de los primeros pasos en este camino es la identificación y clasificación de características de calidad, las cuales se organizan generalmente de forma jerárquica y se documentan en modelos de calidad.

Si bien cada organización de salud debería implementar estos controles, existen razones por las cuales es deseable que una plataforma de integración también los realice, en particular, cuando las organizaciones intercambian datos. En primer lugar, las organizaciones participantes en un sistema de salud a nivel nacional son en general autónomas y heterogéneas, por lo que los criterios de calidad de datos pueden variar de una a otra. En segundo lugar, algunos requerimientos de calidad pueden referir a aspectos de todo el sistema de salud por lo que su control podría requerir datos disponibles sólo en la plataforma (p. ej. intercambios de datos previos). Por último, puede ser conveniente que una plataforma de integración implemente controles de calidad de datos para que las organizaciones puedan delegar parte de estos controles a la misma, generando de esta manera economía de escala.

Si bien las plataformas de integración, en particular las basadas en ESBs, brindan mecanismos que pueden utilizarse para controlar la calidad de los datos que se intercambian a través de ellas, no proveen soluciones nativas para abordar este tema [7], ni de forma general ni para el área de la salud.

Este artículo propone un ESB consciente de la calidad de datos (DQ-ESB) para plataformas de integración en el área de la salud que permite evaluar y hacer cumplir requerimientos de calidad en el intercambio de datos entre organizaciones. DQ-ESB maneja un modelo de calidad de datos extensible con características de calidad que fueron identificadas en el contexto de la salud en Uruguay. La propuesta proporciona además soluciones de calidad de datos basadas en mecanismos nativos de los ESB (p. ej. transformación de datos) y se implementó en base al producto Switchyard.

El resto del artículo se organiza de la siguiente manera. En la Sección II se presenta marco teórico relevante para la

propuesta. En la Sección III se presenta un modelo de calidad de datos para salud aplicable al contexto de la plataforma de salud uruguaya. En la Sección IV se propone un ESB consciente de la calidad de datos para plataformas de salud. En la Sección V se describen pruebas realizadas y un ejemplo de uso. Por último, en la Sección VI se analiza trabajo relacionado y en la Sección VII se presentan conclusiones.

## II. MARCO TEÓRICO

Esta sección presenta marco teórico relevante para la propuesta de este trabajo.

### A. Calidad de Datos en el Área Salud

La información sanitaria comprende datos organizados recogidos sobre un paciente individual, o el resumen de información acerca de la experiencia del paciente con el proveedor de servicios de salud [8]. Es la representación formal de un evento clínico o una característica de un paciente [6]. Esta información tiene características especiales dada su sensibilidad y la importancia de contar con información precisa, oportuna y accesible para la correcta atención sanitaria de la población.

La calidad de esta información juega un papel importante en la toma de decisiones y definición de políticas públicas, en la elaboración de estadísticas, la vigilancia epidemiológica, la planificación de servicios y distribución de recursos. Para los prestadores de salud, la información con buena calidad, ayuda en la coordinación de la atención entre los servicios prestados y mejora la calidad, gestión y eficiencia de los mismos, impactando en la gestión del presupuesto para el mantenimiento de los servicios [6].

Son varios los factores que impactan la calidad de los datos que generan la información sanitaria, por ejemplo, no utilizar estándares y la falta de uso de un vocabulario común que permita que los datos sean legibles, compartidos e interpretados por todos los actores. Otro motivo que repercute en la calidad de los datos son los errores en los diseños de los formularios (p. ej. en papel o electrónicos) que sirven como herramienta para recolectarlos, falta de controles en la entrada de los datos y falta de entrenamiento al personal, entre otros.

### B. Modelos de Calidad de Datos

La identificación y clasificación de características de calidad (p. ej. exactitud sintáctica) es una de las actividades principales cuando se quieren abordar aspectos de calidad de datos. Estas características se organizan generalmente de forma jerárquica y se documentan en modelos de calidad.

La forma en que se organizan los modelos de calidad de datos varía dependiendo de la propuesta. En particular, en [9] se propone un metamodelo de calidad que define una organización en base a dimensiones, factores, métricas y métodos. Una dimensión captura una faceta de calidad a alto nivel (p. ej. exactitud) mientras que un factor representa un aspecto particular de una dimensión (p. ej. exactitud sintáctica). Una métrica es un instrumento utilizado para medir un factor de calidad (p. ej. formato fecha) y un método es un proceso que implementa una determinada métrica (p. ej. un programa que comprueba si una fecha tiene un formato determinado).

### C. Enterprise Service Bus

Un Bus de Servicios Empresariales (*Enterprise Service Bus, ESB*) es una plataforma de integración basada en estándares que combina, servicios web, transformación de datos, mensajería y ruteo inteligente para conectar y coordinar, de forma confiable, la interacción de aplicaciones diversas con integridad transaccional [3].

Los ESBs brindan una capa de procesamiento intermedia que provee lógica de integración y comunicación reutilizable, para posibilitar la interacción entre clientes y servicios [3]. Para esto, los ESBs aceptan pedidos en forma de mensajes sobre los cuales se pueden realizar operaciones de mediación para solucionar heterogeneidades entre clientes y servicios.

Si bien las funcionalidades que ofrecen los ESB pueden variar de acuerdo al proveedor, en general, todos brindan capacidades de conectividad y adaptadores, transformación de mensajes, ruteo inteligente, mecanismos para la implementación de flujos de mediación, mensajería asincrónica y capacidades de monitoreo y administración.

La transformación de mensajes posibilita, por ejemplo, que aplicaciones que usan distintos formatos o modelos de datos puedan comunicarse [3]. La mayoría de los ESBs permiten especificar y ejecutar transformaciones con el estándar XSLT<sup>1</sup>.

El ruteo inteligente es la capacidad por la cual el ESB determina, en tiempo de ejecución, el destino de un mensaje de acuerdo a distintos factores. Algunos tipos de ruteo inteligente son el ruteo basado en contenido y el ruteo basado en itinerario [3]. En particular, el ruteo basado en contenido determina el destino del mensaje en base a su contenido (p. ej. un valor en el cabecal del mensaje, un valor en los datos de negocio) [3].

Los flujos de mediación permiten especificar una secuencia de operaciones de mediación (p. ej. transformación, ruteo) a ejecutar en los mensajes. Estos flujos están pensados para procesos simples y de corta duración, abordando principalmente problemáticas de integración y comunicación [10]. Varios productos ESB permiten especificar flujos de mediación en base a los Patrones de Integración Empresarial (*Enterprise Integration Patterns, EIP*) [11].

### D. Estándares para el Área de Salud

La estandarización en el área de la salud surge por la necesidad de compartir información clínica entre diversas organizaciones, representada de forma heterogénea. A partir del uso de estándares se pretende homogeneizar la información o la interpretación de la misma de forma tal que pueda ser entendida e interpretada sin errores. Algunas organizaciones que están trabajando en la creación y difusión de estándares son *Health Level Seven (HL7) International* [12] e *International Health Terminology Standards Development Organization (IHTSDO)* [13].

En particular, HL7 desarrolló los estándares HL7 [12] y *Clinical Document Architecture (CDA)* [14]. HL7 es un estándar de mensajería para el intercambio de información electrónica en el dominio clínico, mientras que CDA provee

<sup>1</sup><https://www.w3.org/TR/xslt>

un modelo para el intercambio de documentos clínicos. Este último permite que los documentos sean factibles de ser leídos por una máquina y puedan ser recuperados, así como utilizados de forma sencilla.

Por su parte, IHTSDO trabaja en un diccionario que provee códigos, términos, sinónimos y definiciones utilizados en la documentación clínica y en reportes: *Systematized Nomenclature of Medicine Clinical Terms* (SNOMED CT) [13]. SNOMED es una colección de términos clínicos sistemáticamente organizados y procesables por computadora.

Por último, el estándar de Clasificación Internacional de Enfermedades (CIE) [15] fue propuesto por la Organización Mundial de la Salud, para la codificación de la información de causas de morbilidad y mortalidad y con el fin de poder comparar valores a nivel nacional e internacional. La CIE se ha convertido en el estándar internacional para la gestión de la salud y estudios epidemiológicos. Actualmente, se encuentra en la versión 11.

### III. MODELO DE CALIDAD DE DATOS PARA SALUD

En esta sección se describe la Plataforma de Salud de Uruguay y el diseño del Modelo de Calidad de datos propuesto para dicho contexto. En particular, el diseño del modelo constó de tres fases: i) identificación de características de calidad, ii) organización de las características en dimensiones y factores, y iii) definición de métricas para los factores.

Cabe destacar que para el modelo propuesto, se tomaron en cuenta características que se consideraron relevantes y representativas para el alcance del trabajo, dejando para un trabajo futuro un relevamiento más exhaustivo.

#### A. Plataforma de Salud de Uruguay

El desarrollo de la Historia Clínica Electrónica Nacional (HCEN) en Uruguay se propuso como uno de los objetivos de la Agenda Digital 2011 - 2015 [16]. Con este fin, surgió el Programa Salud.uy, como un acuerdo entre la Agencia para el Gobierno Electrónico y la Sociedad de la Información y Conocimiento (AGESIC) [17], Presidencia de la República, el Ministerio de Economía y Finanzas, y el Ministerio de Salud Pública. Salud.uy tiene como objetivo proponer soluciones para una fluida y sistemática integración de servicios de salud, en particular, para el intercambio de información clínica, entre las instituciones participantes del Sistema Nacional Integrado de Salud (SNIS). Salud.uy se enfoca en el desarrollo de la HCEN y en la creación de un Banco Nacional de Registros de Salud [5].

Para dar soporte al desarrollo de la HCEN, AGESIC y Salud.uy están trabajando en la construcción de una Plataforma de Salud que, apoyándose en la Plataforma de Interoperabilidad (PDI) [4] puesta en marcha por AGESIC, aborde problemáticas concretas al área de la salud. Esta plataforma debe permitir intercambiar información clínica, entre los diferentes actores, de forma eficiente y efectiva así como brindar servicios específicos para el área. La Figura 1 presenta de forma simplificada la Plataforma de Salud de Uruguay.

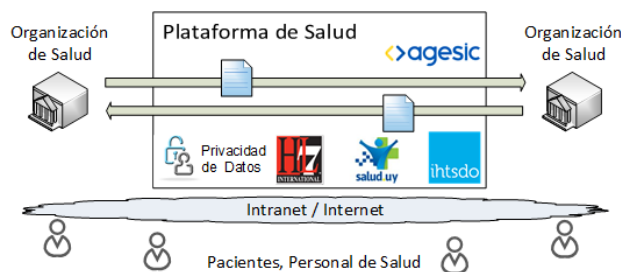


Figura 1: Plataforma de Salud de Uruguay.

Las organizaciones de salud intercambian información sanitaria a través de la plataforma mediante la tecnología de servicios web. En particular, el intercambio de información se realiza mediante mensajes XML siguiendo los estándares SOAP, WS-Security, HL7 y CDA, entre otros.

Este intercambio de mensajes debe cumplir con requerimientos de conformidad, en particular en el área de calidad de datos, que pueden surgir de leyes, estándares y lineamientos establecidos por Salud.uy y AGESIC, entre otros. Estos requerimientos pueden aplicar a todo el mensaje (p. ej. el mensaje debe estar firmado digitalmente), a determinado elemento del mensaje (p. ej. el valor del elemento debe ser un término de SNOMED) o a un conjunto de elementos del mensaje (p. ej. el valor del elemento ciudad debe corresponder a una ciudad del país especificado en el elemento país).

#### B. Identificación de Características de Calidad

Para la identificación de características de calidad de datos relevantes para el contexto descrito en la Sección III-A se tomaron como base tres tipos de fuentes: fuentes internacionales, fuentes nacionales y trabajos académicos en el área de calidad de datos. La Figura 2 muestra gráficamente las distintas fuentes que fueron consultadas.



Figura 2: Identificación de Características de Calidad.

A partir de las fuentes nacionales e internacionales se obtuvieron requerimientos de calidad deseables para los datos que se intercambian entre las diferentes organizaciones. De estos requerimientos y del análisis de diversos trabajos en el área de calidad de datos, se obtuvo el grupo de características a considerar para el diseño del Modelo de Calidad.

Entre las fuentes internacionales se consultaron HL7, IHTSDO, la Organización Mundial de la Salud y el estándar ISO/IEC 25012:2008 [18]. En particular, de estas fuentes se

obtuvieron requerimientos que establecen que los mensajes tienen que respetar estándares como HL7 y CDA, así como el uso de codificación CIE y terminología SNOMED. Por otro lado, a partir del estándar ISO/IEC 25012:2008, se identificaron características relevantes para el área. Algunas características identificadas son Conformidad y Exactitud Sintáctica.

Por otro lado, entre las fuentes nacionales se consultaron a AGESIC, leyes vigentes en Uruguay, el Programa Salud.uy [19] y el Ministerio de Salud Pública (MSP) [20]. De estas fuentes, se obtuvieron requerimientos que establecen que los datos deben adoptar el uso de Modelos de Referencia (Modelo de Referencia de Personas y Direcciones [21]) y la implementación de leyes nacionales (p. ej. Ley N° 18.600 de Documento Electrónico y Firma Electrónica [22]). Por otro lado, se identificó como requerimiento el uso del estándar CIE10 (versión 10 del CIE) para la codificación de la información de causas de morbilidad y mortalidad y el uso de los datos mínimos establecidos a nivel nacional para el envío de información de eventos específicos. De estos requerimientos surgen, p. ej., la característica Exactitud Sintáctica y características en el área de la Completitud y Seguridad.

Por último, se consultaron y analizaron diversos trabajos que buscaron identificar distintas características de calidad deseables sobre datos en el área de salud [23] [24] [6] [25]. Del análisis se pudo concluir que no hay acuerdo sobre el conjunto de características relevantes a considerar en el área. Esto se debe a que los autores utilizan distintas metodologías para obtener dicho conjunto, e incluso, utilizan el mismo nombre para representar cosas semánticamente diferentes, así como diferentes nombres para representar la misma característica. Algunas de las características obtenidas de este análisis son Integridad Relacional del área Consistencia y la característica Oportunidad del área Frescura.

Luego de analizados los tres tipos de fuentes, se obtuvo un grupo consolidado de características de Calidad de Datos en el área de salud que resulta de interés para el alcance de este trabajo. Esta selección se basó en el uso repetido de la característica a lo largo de las diversas fuentes, la posibilidad de medir la calidad del dato durante el intercambio del mismo y la relevancia de la característica para el presente trabajo.

### C. Organización de Características

Luego de obtener las características de calidad relevantes para el contexto de trabajo presentado, las mismas se organizaron en un Modelo de Calidad estructurado en base a los conceptos presentados en la Sección II-B.

El Modelo de Calidad está compuesto por las Dimensiones y Factores que se presentan en la Tabla I. La descripción de cada Factor resulta del análisis de las definiciones en las distintas fuentes, tomando como resultado la que se consideró más apropiada para el contexto.

### D. Definición de Métricas

Como último paso en el diseño del modelo, se definieron métricas para los factores de calidad indicando para cada una

Tabla I: Dimensiones de Calidad de Datos

Dimensión	Factores	Descripción	Fuente
Completitud	Nulos	Los datos requeridos deben estar presentes.	[21] [19] [6] [25] [18]
	Datos Mínimos	Los datos mínimos definidos para el evento deben estar presentes.	[21] [19] [6] [25]
Frescura	Oportunidad	Los datos deben ser oportunos para su fin.	[24] [6] [25]
	Actualidad	Los datos deben ser actuales.	[25] [6] [18]
Seguridad	Credibilidad	Los datos deben estar firmados para poder identificar el autor.	[22] [23] [18]
	Conformidad	Los valores deben respetar el estándar definido.	[13] [14] [12] [15] [6] [25] [18]
Consistencia	Integridad Relacional	Los datos deben ser consistentes entre sí.	[24] [6] [18]
Exactitud	Exactitud Sintáctica	Los valores de los datos deben representar valores válidos dentro de un dominio y reglas establecidas.	[15] [20] [13] [21] [19] [23] [6] [25] [18]
	Exactitud Semántica	Los valores de los datos deben reflejar el valor real.	[24] [25] [6] [18]
	Precisión	Los datos deben estar representados con la precisión establecida.	[25] [6]

de ellas: cómo se mide (i.e. semántica), unidad utilizada (p. ej. Boolean) y el objeto medible al cual aplica (i.e. granularidad).

Dado que se quiere evaluar la calidad de los datos incluidos en los mensajes que se intercambian a través de la plataforma, se definieron tres objetos medibles: el propio mensaje que se intercambia, un elemento dentro del mensaje y un conjunto de elementos. La granularidad de cada métrica se especifica entonces en base a estos objetos.

La Tabla II presenta el conjunto de métricas que se consideraron relevantes para el trabajo y que apunta a ser descriptivo de las métricas que podría incluir el modelo.

## IV. ESB PARA CALIDAD DE DATOS

Esta sección describe DQ-ESB: un ESB que permite evaluar y hacer cumplir requerimientos de calidad de datos en los intercambios de información entre organizaciones de salud.

### A. Descripción General

La idea general de la propuesta es extender las capacidades que ofrece un ESB, para que sea posible el monitoreo, control y aseguramiento de la calidad de datos en los mensajes que las organizaciones intercambian a través de él. En la Figura 3 se puede observar cómo la solución de Calidad de Datos se incorpora a una Plataforma de Salud basada en un ESB.

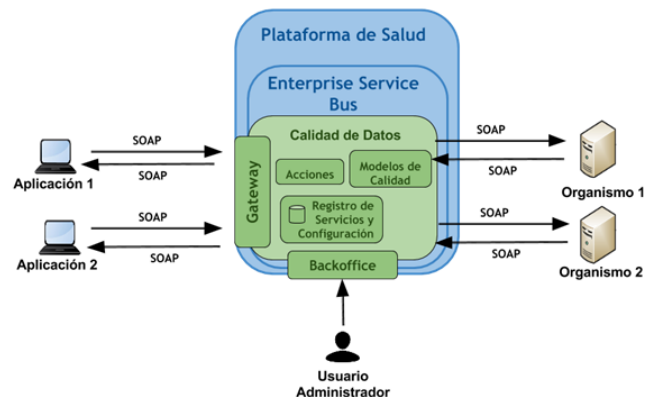


Figura 3: ESB con Componente de Calidad.

Los principales componentes de la solución propuesta son:

Tabla II: Métricas

Dimensión	Factor	Métrica	Semántica	Unidad	Granularidad
Exactitud	Exactitud Sintáctica	FormatoFecha1()	¿El dato cumple el formato yyyy-mm-dd?	Boolean	Elemento
		DominioSexo()	¿El dato se encuentra dentro del dominio definido (1, 2 y 9) que indican Masculino, Femenino, No Aplica respectivamente?	Boolean	Elemento
		DominioPaisISOAlpha3()	¿El dato se encuentra dentro del dominio definido por la ISO 3166-1?	Boolean	Elemento
	Exactitud Semántica	FechaAnteriorHoy()	¿La Fecha es menor o igual a la de hoy?	Boolean	Elemento
		FechaPosteriorHoy()	¿La Fecha es igual o mayor a la fecha de hoy?	Boolean	Elemento
		FechaDiferenciaHoy()	Resta la fecha a comparar con la fecha de hoy	Día	Elemento
	Precisión	PrecisionDecimales()	Calcula la cantidad de decimales	Entero	Elemento
Consistencia	Integridad Relacional	IntegridadDepartamentoLocalidad()	¿La Localidad del mensaje corresponde a una Localidad del Departamento relacionado?	Boolean	Conjunto de elementos (Departamento, Localidad)
		IntegridadEnfermedadSexo()	¿La Enfermedad indicada en el mensaje es consistente con el Sexo?	Boolean	Conjunto de elementos (Enfermedad, Sexo)
Frescura	Actualidad	EdadDato()	Calcula la diferencia entre la fecha del dato en el mensaje y la fecha de hoy	Día	Elemento
	Oportunidad	EdadDato()	Calcula la diferencia entre la fecha del dato en el mensaje y la fecha de hoy. Es la misma métrica usada para el factor Actualidad.	Día	Elemento
Interoperabilidad	Conformidad	ExisteCIE10()	¿El elemento representa un código CIE10 válido?	Boolean	Elemento
		ExisteSNOMED()	¿El elemento representa un término SNOMED válido?	Boolean	Elemento
		ElementoHL7()	¿El mensaje contiene elementos HL7 validos?	Boolean	Mensaje
		EstructuraCDA()	¿El mensaje tiene la estructura de un documento CDA?	Boolean	Mensaje
Complejidad	Datos Mínimos	DatosMínimosEmergencia()	Para un mensaje que describe un evento de Emergencia, ¿se contemplan todos los datos mínimos definidos por el Programa Salud.uy?	Boolean	Elemento
		Nulo()	¿El dato es igual al valor nulo?	Boolean	Elemento
	Nulos	VacioEntero()	¿El dato es igual al valor 0?	Boolean	Elemento
		VacioTexto()	¿El dato es igual al valor ?	Boolean	Elemento
Seguridad	Credibilidad	MensajeFirmado()	¿El mensaje está firmado digitalmente?	Boolean	Mensaje
		ElementoFirmado()	Para un elemento particular, ¿está firmado digitalmente?	Boolean	Elemento

- **Gateway:** Componente encargado de centralizar el acceso de los clientes a los servicios brindados por las organizaciones a través de la plataforma, así como también de enviarles las respuestas.
- **Backoffice:** Componente web para gestionar los organismos que se integran a través de la plataforma y de los servicios web que estos publican. También permite gestionar los Modelos de Calidad de Datos en la plataforma así como configurar Acciones a realizar cuando se detecta que un requerimiento de calidad de datos no se cumple.
- **Registro de Servicios y Configuración:** Componente que almacena datos de configuración y metadatos de los servicios que se publican a través de la plataforma.
- **Modelos de Calidad:** Componente que implementa los Métodos de medición para las métricas definidas en los modelos de calidad que la plataforma soporta.
- **Acciones:** Componentes que realizan acciones en caso que la evaluación de la calidad de datos no sea satisfactoria. Estas Acciones pueden ser notificar a un administrador, cancelar el mensaje respondiendo al cliente con el

error correspondiente o ejecutar mecanismos para tratar de asegurar la calidad (p. ej. transformación del mensaje).

La Figura 4 presenta un ejemplo en el cual se describe el funcionamiento general de la plataforma. En este ejemplo, el organismo Dirección General de Registro de Estado Civil (DGREC) invoca la operación obtenerCertificado del Ministerio de Salud Pública (MSP), la cual retorna un certificado de defunción. Existe además un Modelo de Calidad donde se especifica un Método de Medición que determina si el formato del elemento fechaDefuncion, parámetro de entrada de la operación mencionada, tiene el formato yyyy-mm-dd. Como en este caso el mensaje no cumple con este control, se ejecutan las Acciones asociadas al Método: se envía un correo electrónico notificando el incumplimiento de dicho control de calidad y se realiza una transformación XSLT al mensaje, para convertir la fecha al formato esperado.

De forma detallada, el flujo completo es el siguiente: la DGREC envía un pedido a la plataforma para invocar al servicio del MSP (1). En el *Gateway* se realizan validaciones y el registro del mensaje de entrada. En este caso, según la

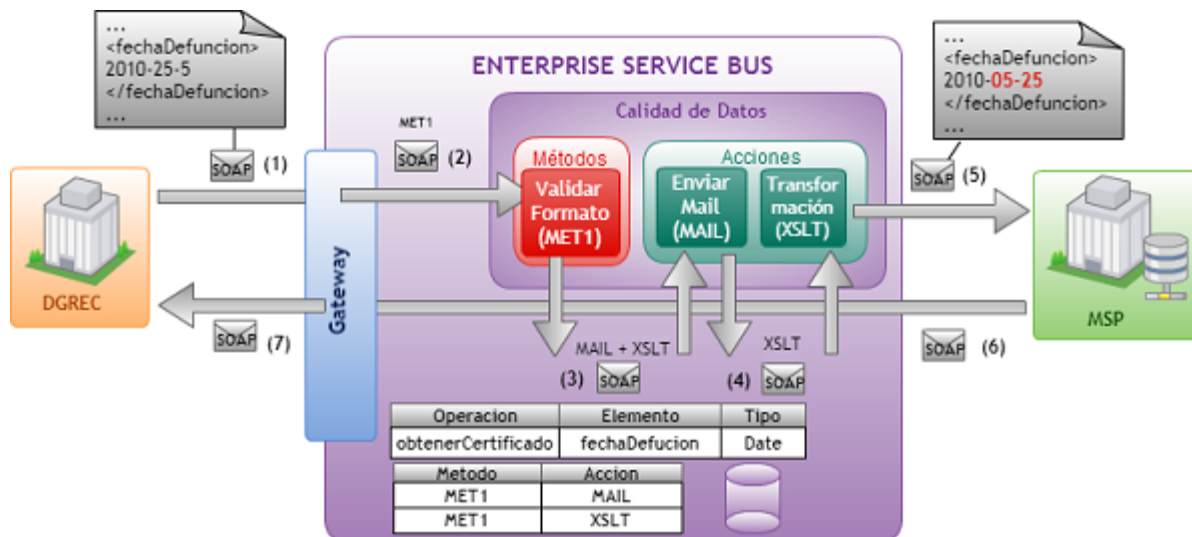


Figura 4: Ejemplo de Control de Calidad de Datos en ESB.

operación y sus elementos se detecta que debe evaluarse el método MET1, por lo que se redirige el mensaje hacia el Método correspondiente (2). Se ejecuta entonces el Método y, al no cumplir con el requerimiento del formato, el mensaje es redirigido hacia la primera de las Acciones asociadas al mismo, que en este caso se corresponde con el envío de una notificación por correo electrónico (3). Luego el pedido se envía hacia el componente encargado de ejecutar la siguiente Acción, donde se lleva a cabo una transformación XSLT (4). El mensaje final es enviado hacia el servicio que el cliente original desea consumir (5). Una vez que el MSP termina de procesar el pedido, retorna la respuesta al ESB (6), y al no existir Métodos configurados para los elementos del mensaje de respuesta, es enviada hacia la DGREC por medio del Gateway (7).

### B. Modelo Conceptual

La Figura 5 presenta el modelo conceptual de la solución propuesta.

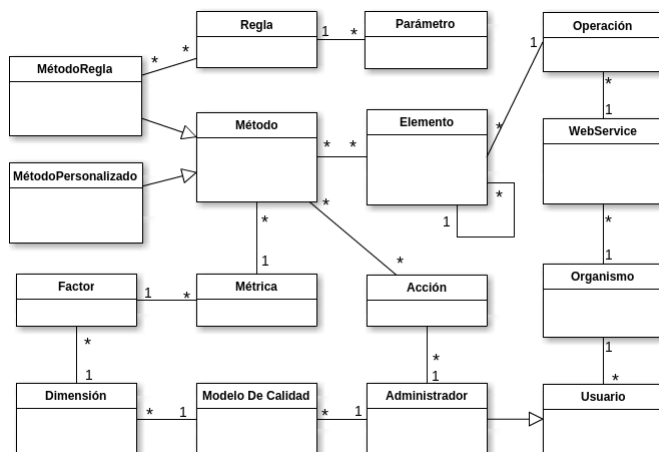


Figura 5: Modelo Conceptual.

En particular los organismos publican servicios web que pueden tener varias operaciones. Cada operación tiene un conjunto de elementos de entrada y otro de salida. Cada elemento puede asociarse con métodos de evaluación que retornan verdadero o falso de acuerdo a criterios que se especifican en la definición del método. Si el método retorna un valor falso significa que el elemento no cumple con algún requerimiento de calidad de datos.

El sistema permite la especificación de métodos de evaluación de dos formas: definiendo Reglas o utilizando Métodos personalizados. La especificación de métodos mediante reglas permite a un usuario sin conocimientos de programación implementar Métodos por medio de la combinación de reglas de comparación, funciones de conjuntos, operaciones lógicas y/o expresiones regulares. Por ejemplo, un usuario puede definir un conjunto de Reglas sobre los elementos del mensaje y determinar la relación lógica que éstas deben cumplir para retornar un valor verdadero. A modo de ejemplo, en la Tabla III se presentan dos Reglas: la primera especifica que el valor del Parámetro Sexo tiene que estar contenido en la lista "1, 2, 9" (Masculino, Femenino, No aplica) y la segunda especifica que el Parámetro Documento tiene que tener un valor de largo menor que 9. Por otro lado, la Tabla IV especifica un método de calidad para la métrica con identificador 2 el cual retorna verdadero cuando el valor del elemento Sexo está contenido en la lista "1, 2, 9" y el largo del valor del elemento Documento es mayor o igual a 9. En la definición del método también se especifica que de no cumplirse con el requerimiento se tiene que ejecutar la Acción con identificador 3 (p. ej. enviar un correo electrónico).

Tabla III: Ejemplos de Reglas

Id Regla	Operación	Parámetro	Valor	Método
R1	Contenido	Sexo	1, 2, 9	1
R2	LargoMenor	Documento	9	1



Tabla IV: Ejemplo de Método de Calidad

Rel. entre reglas	Métrica	Acción
R1 and (not R2)	2	3

Por otra parte, usuarios técnicos pueden desplegar en el sistema implementaciones personalizadas, brindando más flexibilidad y permitiendo crear Métodos de mayor complejidad.

### C. Detalles de Implementación

Esta sección presenta detalles del desarrollo del prototipo, del cual se dejó disponible una demostración<sup>2</sup>. La implementación del prototipo se realizó en JavaEE 7 utilizando la plataforma JBoss EAP 6.4. En la Figura 6 se muestra un diagrama de los componentes implementados y las principales tecnologías utilizadas para su implementación.

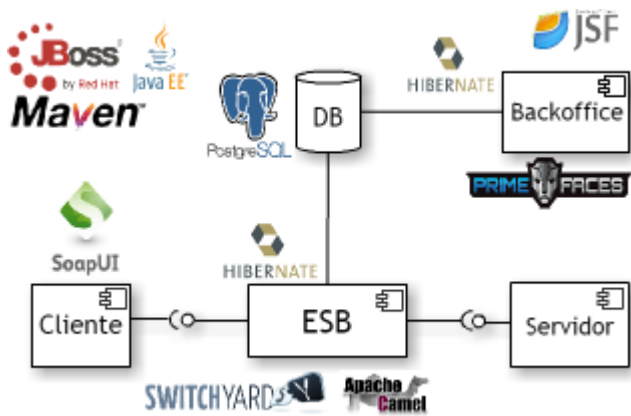


Figura 6: Tecnologías Utilizadas.

Apache Maven se utilizó para definir la estructura modular del proyecto, compilación y gestionar las dependencias a librerías utilizadas. PostgreSQL se utilizó para almacenar todas las estructuras gestionadas por el Backoffice, como Modelos de Calidad, organismos, servicios web y también el registro de logs. Hibernate se utilizó para la correspondencia objeto-relacional de forma de facilitar el almacenamiento y consultas de los datos que maneja la plataforma. Apache Camel es un framework Java de código abierto, enfocado en facilitar la integración de aplicaciones sobre distintas plataformas. Para ello provee implementaciones de los EIP más utilizados. En particular, este framework se utilizó para resolver el ruteo de los mensajes en la plataforma. Switchyard tiene todas las características esperables en un típico ESB y está integrado con Apache Camel añadiendo, de esta forma, funcionalidades de ruteo, transformación de mensajes y orquestación de servicios. Java Server Faces (JSF) es un estándar Java para el desarrollo de interfaces de usuario web del lado del servidor<sup>3</sup>. Primefaces es una librería de código abierto, que cuenta con un conjunto de componentes enriquecidos para JSF. JSF y Primefaces se utilizaron para la implementación del Backoffice.

<sup>2</sup><https://www.fing.edu.uy/inco/grupos/lins/demos/dq-esb.mp4>

<sup>3</sup><http://www.oracle.com/technetwork/java/javaee/javaserverfaces-139869.html>

## V. EJEMPLO DE USO Y PRUEBAS

En esta sección se describe un ejemplo de uso y se presentan resultados de pruebas realizadas.

### A. Ejemplo de Uso

El ejemplo de uso está basado en el servicio Consulta Certificado de Defunción Electrónico<sup>4</sup> provisto por el MSP y describe cómo son modificados mensajes que recibe la plataforma, con el fin de asegurar el cumplimiento de los requerimientos de calidad establecidos.

Una de las operaciones del servicio es obtenerCertificado que recibe como parámetros de entrada numeroCertificado, tipoDocumento, paisDocumento y numeroDocumento. Para el ejemplo de uso se considera que hay definido un requerimiento de calidad para el elemento paisDocumento en base a la métrica DominioPaisISOAlpha3 del factor Exactitud Sintáctica definido en el modelo presentado en la Sección III. El requerimiento de calidad establece que el valor de esa métrica debe ser true para dicho elemento (i.e. el valor del elemento paisDocumento debe estar en la lista de países definidos en ISO 3166-1). También se considera que hay definida una acción a ejecutar en caso de que ese requerimiento de calidad no se cumpla: una transformación XSLT que cambia valores no válidos del elemento paisDocumento por valores válidos de acuerdo a ISO 3166-1.

La Figura 7 presenta el cuerpo del mensaje SOAP que recibe la plataforma para invocar a dicha operación.

```
<soapenv:Body>
  <con:obtenerCertificado>
    <numeroCertificado>1000</numeroCertificado>
    <tipoDocumento>68909</tipoDocumento>
    <paisDocumento>URUGUAY</paisDocumento>
    <numeroDocumento>44072453</numeroDocumento>
  </con:obtenerCertificado>
</soapenv:Body>
```

Figura 7: Mensaje SOAP - Request.

Dado que el mensaje no cumple con el requerimiento de calidad, se ejecuta la transformación que se presenta en la Figura 8 para que el valor del elemento cumpla con el requerimiento establecido antes de que se invoque al servicio.

```
<xsl:stylesheet version="1.0" .....>
  <xsl:template match="node()|@"*>
    <xsl:copy>
      <xsl:apply-templates select="node()|@"*/>
    </xsl:copy>
  </xsl:template>
  <xsl:template match="paisDocumento/text()[.='URUGUAY']">
    URY </xsl:template>
</xsl:stylesheet>
```

Figura 8: Transformación XSLT.

La Figura 9 presenta el cuerpo del mensaje SOAP luego de la transformación. De esta forma, el servicio del MSP recibe

<sup>4</sup><https://www.agesic.gub.uy/innovaportal/v/5333/9/agesic/consulta-certificado-de-defuncion-electronico.html>

un mensaje de solicitud válido de acuerdo a los requerimientos establecidos.

```
<soapenv:Body>
  <con:obtenerCertificado>
    <numeroCertificado>1000</numeroCertificado>
    <tipoDocumento>68909</tipoDocumento>
    <paisDocumento>URY</paisDocumento>
    <numeroDocumento>44072453</numeroDocumento>
  </con:obtenerCertificado>
</soapenv:Body>
```

Figura 9: Mensaje SOAP - *Request*.

Por último, la Figura 10 presenta el cuerpo del mensaje SOAP que hubiera recibido el cliente en caso de que, en lugar de una transformación, se hubiera configurado que se cancelara el mensaje si no se cumplía con el requerimiento de calidad.

```
<soap:Fault>
  <soap:faultcode>400</soap:faultcode>
  <soap:faultstring>
    No se cumple con requerimiento de calidad.
  </soap:faultstring>
  <soap:detail><error>
    1152 – DominioPaisISOAlpha3: paisDocumento
    no cumple con formato requerido..
  </error></soap:detail>
</soap:Fault>
```

Figura 10: Mensaje SOAP Error - *Response*.

## B. Pruebas No Funcionales

Se realizaron pruebas para comprobar el *overhead* que introduce la solución propuesta en el tiempo de respuesta de los servicios. En particular, se ejecutaron cinco tipos de invocaciones a un servicio: i) invocación directa (i.e. sin pasar a través de la plataforma), ii) invocación a través de la plataforma pero sin aplicar métodos de evaluación de calidad, iii) invocación a través de la plataforma aplicando métodos de evaluación de la calidad con datos que cumplieran con los requerimientos de calidad, iv) invocación a través de la plataforma aplicando métodos de evaluación con datos en el *request* que no cumplieran con los requerimientos de calidad, y v) invocación a través de la plataforma aplicando métodos de evaluación con datos en el *response* que no cumplieran con los requerimientos de calidad.

Para esto se simularon cinco clientes ejecutando consultas al servicio web de forma concurrente, en intervalos de 1000ms durante 60 segundos. Se registraron para cada una de las estrategias empleadas los tiempos de procesamiento, los cuales fueron utilizados para realizar una comparación con la invocación directa al servicio. Los tiempos promedio y el *overhead* generado por la solución en cada caso, se pueden consultar en la Tabla V.

De los resultados obtenidos se puede concluir que el solo hecho de pasar por la plataforma, como es el caso de la estrategia 2, agrega un *overhead* aproximado de 118ms. Esto puede explicarse debido a que una vez que un mensaje ingresa

Tabla V: Overhead Generado según Estrategias

ID	Estrategia	Promedio (ms)	Overhead (ms)
1	Invocación Directa	7	-
2	Sin Aplicación Métodos Evaluación	125	118
3	Con Métodos (request no cumple con calidad)	132	125
4	Con Métodos (datos cumplen con calidad)	256	249
5	Con Métodos (response no cumple con calidad)	268	261

a la plataforma, se agrega un tiempo fijo de procesamiento para realizar operaciones de validación del formato y de la validez del mensaje de acuerdo a las configuraciones de la plataforma. Luego puede observarse una diferencia entre las estrategias sin métodos (i.e. estrategia 2) y con problemas de calidad en el *request* (i.e. estrategia 3), con respecto a las que no tienen problemas de calidad (i.e. estrategia 4) o estos son detectados en el *response* (i.e. estrategia 5). Esto puede explicarse debido a que para esta prueba, en la estrategia 2 no se evalúan Métodos y en la 3 se evalúan sólo siete métodos, que son los aplicados a un mensaje *request* que consume ese servicio. Mientras que para las estrategias 4 y 5 se evalúan 34 Métodos que aplican a los elementos del *request* o a los del *response*, además de agregarse la invocación al servicio final.

Por otro lado, en la Figura 11 se puede apreciar gran similitud en cuanto a los tiempos de respuesta de los mensajes no cancelados y los cancelados en el *response*. Esto permite apreciar que el hecho de cancelar un mensaje prácticamente no implica *overhead* ante igual cantidad de Métodos evaluados.

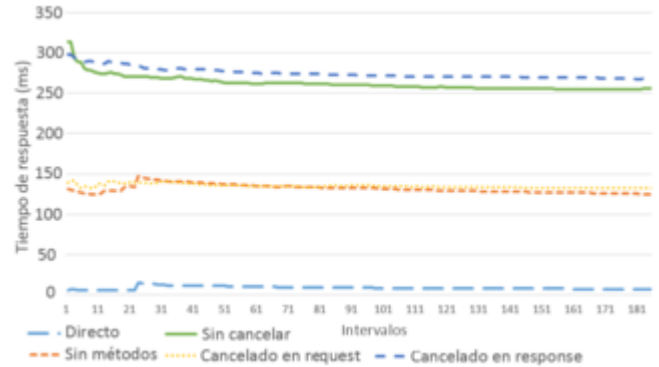


Figura 11: *Overhead* Generado.

Al ejecutar las pruebas, se pueden diferenciar dos claros flujos por los cuales puede pasar un mensaje en la solución: el monitoreo y la validación. El flujo de monitoreo no agrega ninguna lógica sobre el mensaje que llega a la solución, más que realizar un registro y realizar el ruteo del mismo hacia el destino pertinente. Por otro lado, la validación requiere de la ejecución de diversos cálculos y evaluaciones, que determinan qué Métodos y Acciones deben ser invocados. Se puede observar que se genera un crecimiento con un exponente tan bajo que se necesita un gran número de Métodos/Acciones (más de 30 para un mensaje) para que el tiempo de respuesta supere la franja de lo lineal. El *overhead* es generado mayormente por el hecho de tener que ejecutar Métodos y Acciones, más que por la cantidad de los mismos, al involucrar los componentes que se encargan de realizar dichas tareas en el flujo del mensaje.



Cabe mencionar que parte del *overhead* podría ser disminuido sensiblemente (hasta en un 20 %) utilizando componentes Switchyard fijos, y no invocándolos dinámicamente según el contexto o el contenido del mensaje. De todas formas, implementando una invocación dinámica de los mismos da la flexibilidad de introducir nuevos componentes de forma más ágil que si estas invocaciones no pudieran ser configurables.

Las pruebas realizadas permiten concluir que la solución propuesta no introduce un *overhead* que pueda introducir limitaciones importantes en su utilización.

## VI. TRABAJO RELACIONADO

Esta sección presenta trabajo relacionado a la propuesta.

En primer lugar, algunos trabajos plantean la problemática de calidad de datos en el contexto de un ESB [7][26] y proponen soluciones para abordar este tema apoyándose en este tipo de infraestructura. Por ejemplo, en [7] los autores plantean que cuando hay muchas fuentes de datos con información de calidad, el manejo de la calidad de los datos a nivel de usuario es un desafío. Por esto consideran que, para abordar este problema, el proceso de calidad de datos debe moverse desde el nivel de usuario hasta el servidor de integración. A diferencia con nuestra propuesta, el trabajo no aborda requerimientos de calidad de datos en interacciones inter-organizacionales y se enfoca principalmente en extender un ESB para brindar soluciones que permitan la selección de fuentes de datos de acuerdo a criterios de calidad ante pedidos de datos de usuarios. El trabajo tampoco aborda temáticas de calidad de datos particulares del área de la salud.

Por otro lado, la problemática de calidad de datos en el área de la salud es un tema recurrente de investigación [27][28][29][30], en particular, ante la necesidad de interconectar sistemas heterogéneos para poder compartir información sanitaria. En [29] los autores analizan la calidad de los datos presentes en mensajes HL7 y concluyen que la norma HL7 no siempre es correctamente utilizada, ya que el uso de estructuras personalizadas para transportar datos genera inconsistencias a la hora de integrar dichos datos. En el trabajo se proponen también mecanismos para monitorear la calidad de datos en este tipo de mensajes. En [28], los autores consideran los datos de un entorno *pervasive* de Salud, donde las fuentes de datos incluyen dispositivos inalámbricos que permiten sensar datos de los pacientes. Según los autores, un dato, dentro de un entorno *pervasive*, puede ser visto desde muchas perspectivas que incluyen dimensiones de calidad (p. ej. completitud, frescura, credibilidad) y su contexto. Si bien estos trabajos, al igual que otros en el área [6][31], abordan temas de calidad de datos en un contexto de salud, ninguno de ellos propone un modelo de calidad y tampoco presentan propuestas para monitorear y hacer cumplir requerimientos de calidad de datos en intercambios de información inter-organizacional, en particular, utilizando un ESB.

Hay también varias propuestas que plantean el uso de plataformas de *middleware* para abordar temas de calidad de datos. En [30] se presenta un contexto de aplicación similar al de nuestra propuesta (i.e. organizaciones que intercambian

datos de salud) en el que se utiliza una plataforma de salud inteligente a través de la cual se comunican las organizaciones. La plataforma se basa en una Arquitectura Orientada a Servicios (*Service-Oriented Architecture*, SOA), se implementa mediante distintos componentes entre los que se incluye un ESB y aborda además temas de calidad de datos. En [32] se centran en los requerimientos de calidad de los usuarios y mencionan los posibles conflictos que se pueden presentar entre los mismos. Los autores argumentan que dichos conflictos deben ser resueltos de forma automática y para esto proponen un *middleware* basado en componentes configurables. Por otro lado, en contextos de Internet de las Cosas (*Internet of Things*, IoT) [33][34] la calidad de datos se considera un tema fundamental. Por ejemplo, en [33] los autores introducen un *framework* dentro de una plataforma de *middleware* distribuida para abordar cuestiones de privacidad y calidad de datos. Por otra parte en [34], los autores consideran que un *middleware* de sistemas distribuidos reconfigurable puede aplicarse en entornos de sistemas heterogéneos, no confiables, dinámicos y de gran escala. En particular, proponen un *middleware* dirigido por políticas para abordar requerimientos legales que incluyen temas de calidad de datos. Por su parte, en [35][36] los autores abordan la gestión de calidad de datos en el intercambio de datos maestros utilizando el nuevo estándar ISO 8000. Estos trabajos proponen el uso de SOA para implementar el control de dicho estándar, en especial las partes 100 a 140 que tratan específicamente sobre el intercambio de datos maestros. A diferencia con nuestra propuesta, en estos trabajos no se propone un modelo de calidad de datos para el área salud y tampoco se utilizan las capacidades de un ESB para monitorear y asegurar el cumplimiento de requerimientos de calidad de datos en los intercambios de información.

## VII. CONCLUSIONES

Este artículo presenta DQ-ESB, una plataforma de integración de servicios basada en ESB especializada en el control de calidad de datos en área de la salud. DQ-ESB permite evaluar y hacer cumplir requerimientos de calidad de datos en el contexto de interconexión entre organizaciones.

Concretamente, DQ-ESB consiste en un ESB extendido con funcionalidades de monitoreo, control y aseguramiento de la calidad de datos aplicando operaciones de mediación (p. ej. transformaciones) sobre los mensajes intercambiados. El Modelo de Calidad de datos subyacente, que incluye Dimensiones, Factores y Métricas para la evaluación de la calidad, se basa en estándares y marcos de trabajo internacionales así como en la normativa uruguaya aplicable a datos de salud. Este artículo presenta una primera versión de dicho modelo que, por razones de limitación de espacio, incluye los aspectos que se consideraron más relevantes para el alcance del trabajo.

Los principales aportes de este trabajo consisten en la propuesta de una plataforma especializada para el control de calidad de datos combinando técnicas de integración de servicios, *middleware* de tipo ESB y un Modelo de Calidad de datos instanciado al área de la salud. Asimismo, este trabajo incluye un ejemplo de uso y pruebas experimentales, en

particular sobre el *overhead* generado al utilizar la plataforma de integración, que contribuyen al estudio de factibilidad técnica del enfoque propuesto. Se apunta a que este tipo de mecanismos sea aplicable en una plataforma de integración especializada al área de la salud, como la de Uruguay.

El trabajo futuro consiste en continuar el desarrollo del Modelo de Calidad de datos para salud. Entre otras, se apunta a incorporar características del nuevo estándar ISO 8000 para calidad de datos. También, se apunta a mejorar aspectos de implementación considerando las pruebas realizadas.

Cabe destacar que el enfoque propuesto no está restringido al Modelo de Calidad de datos, lo cual permite su evolución tanto dentro del área específica de salud como de otras áreas.

Más en general, este trabajo constituye un paso más hacia el desarrollo de mecanismos de gestión de normas en el contexto de plataformas de integración inter-organizacionales.

#### ACKNOWLEDGMENT

Este trabajo fue parcialmente financiado por CSIC, Universidad de la República, Uruguay.

#### REFERENCIAS

- [1] G. J. Kuperman, "Health-information exchange: why are we doing it, and what are we doing?" *Journal of the American Medical Informatics Association*, vol. 18, no. 5, p. 678, 2011.
- [2] L. González and R. Ruggia, "A reference architecture for integration platforms supporting cross-organizational collaboration," in *Proceedings of the 17th International Conference on Information Integration and Web-based Applications & Services*, ser. iiWAS '15. New York, NY, USA: ACM, 2015, pp. 92:1–92:4.
- [3] D. Chappell, *Enterprise Service Bus: Theory in Practice*. O'Reilly Media, Jun. 2004.
- [4] L. González, R. Ruggia, J. Abin, G. Llambías, R. Sosa, B. Rienzi, D. Bello, and F. Álvarez, "A Service-Oriented Integration Platform to Support a Joined-Up E-Government Approach: The Uruguayan Experience," in *Advancing Democracy, Government and Governance*, ser. Lecture Notes in Computer Science, A. Kő, C. Leitner, H. Leitold, and A. Prosser, Eds. Vienna, Austria: Springer Berlin Heidelberg, Jan. 2012, pp. 140–154.
- [5] J. Abin, H. Nemeth, and I. Friedmann, "Systems architecture for a nationwide healthcare system," in *MedInfo*, 2015, pp. 12–16.
- [6] W. H. Organization *et al.*, *Improving data quality: a guide for developing countries*. Manila: WHO Regional Office for the Western Pacific, 2003.
- [7] K. Rasta, T. H. Nguyen, and A. Prinz, "A framework for data quality handling in enterprise service bus," in *Third International Conference on Innovative Computing Technology (INTECH 2013)*, 2013, pp. 491–497.
- [8] N. Davis and M. LaCour, *Introduction to Health Information Technology*.
- [9] L. Etcheverry, V. Peralta, and M. Bouzeghoub, "Qbox-Foundation: a Metadata Platform for Quality Measurement," in *Proceedings of the 4eme Atelier Qualité des données et des Connaissances*, Nice, France, 2008.
- [10] M.-T. Schmidt, B. Hutchison, P. Lambros, and R. Phippen, "The enterprise service bus: Making service-oriented architecture real," *IBM Syst. J.*, vol. 44, no. 4, pp. 781–797, Oct. 2005.
- [11] G. Hohpe and B. Woolf, *Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions*. Addison-Wesley Professional, Oct. 2003.
- [12] Health Level Seven International. (2017, Apr.) Health Level Seven. [Online]. Available: <http://www.hl7.org/>
- [13] International Health Terminology Standards Development Organisation. (2017, Apr.) Systematized Nomenclature of Medicine Clinical Terms. [Online]. Available: <http://www.snomed.org/>
- [14] Health Level Seven International. (2017, Apr.) Clinical Document Architecture. [Online]. Available: [http://www.hl7.org/implement/standards/product\\_brief.cfm?product\\_id=7](http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7)
- [15] World Health Organization. (2017, Apr.) International Classification of Diseases. [Online]. Available: <http://www.who.int/classifications/icd/en/>
- [16] AGESIC. (2017, Apr.) Agencia Digital Uruguay 2011-2015. [Online]. Available: [http://www.agesic.gub.uy/innovaportal/v/1443/1/agesic/mapa\\_de\\_ruta:\\_-agenda\\_digital\\_uruguay\\_2011-2015.html?menuderecho=11](http://www.agesic.gub.uy/innovaportal/v/1443/1/agesic/mapa_de_ruta:_-agenda_digital_uruguay_2011-2015.html?menuderecho=11)
- [17] AGESIC. (2017, Apr.). [Online]. Available: <https://www.agesic.gub.uy/>
- [18] ISO/IEC 25012:2008. (2017, Jul.). [Online]. Available: <https://www.iso.org/home.html>
- [19] AGESIC. (2017, Apr.) Programa Salud.uy. [Online]. Available: [http://www.agesic.gub.uy/innovaportal/v/4422/19/agesic/que\\_es.html](http://www.agesic.gub.uy/innovaportal/v/4422/19/agesic/que_es.html)
- [20] Ministerio de Salud del Uruguay. (2017, Apr.). [Online]. Available: <http://www.msp.gub.uy/>
- [21] AGESIC. (2017, Apr.) Modelo de Referencia de Personas y Direcciones. [Online]. Available: <https://www.agesic.gub.uy/innovaportal/v/204/1/agesic/buenas-practicas.html?currentpage=1>
- [22] Poder Legislativo del Uruguay. (2017, Apr.) Ley n° 18.600 Documento Electrónico y Firma Electrónica. [Online]. Available: <https://legislativo.parlamento.gub.uy/temporales/leytemp3217058.htm>
- [23] I. Ibnualim and S. H. Supangkat, "Design of health social media to improve the quality of patient's recovery," in *2012 International Conference on Cloud Computing and Social Networking (ICCCSN)*. IEEE, 2012, pp. 1 – 4.
- [24] M. Kahn, M. Raebel, J. Glanz, K. Riedlinger, and J. Steiner, "A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research," 2012.
- [25] Canadian Institute for Health Information. (2017, Apr.). [Online]. Available: <https://www.cihi.ca/en>
- [26] P. C. Brebner, "Real-world performance modelling of enterprise service oriented architectures: delivering business value with complexity and constraints," *SIGSOFT Softw. Eng. Notes*, vol. 36, no. 5, pp. 85–96, Sep. 2011. [Online]. Available: <http://doi.acm.org/10.1145/1958746.1958762>
- [27] K. Sartipi, M. H. Yarmand, and D. G. Down, "Mined-knowledge and decision support services in electronic health," in *Systems Development in SOA Environments, 2007. SDSOA '07: ICSE Workshops 2007. International Workshop on*, 2007, pp. 10–10.
- [28] J. O'Donoghue and J. Herbert, "Data management within mhealth environments: Patient sensors, mobile devices, and databases," *J. Data and Information Quality*, vol. 4, no. 1, pp. 5:1–5:20, Oct. 2012. [Online]. Available: <http://doi.acm.org/10.1145/2378016.2378021>
- [29] R. J. T. Ferreira, M. E. C. D. Correia, F. N. R. Gonçalves, and R. J. C. Correia, "Data quality in hl7 messages – a real case analysis," in *2015 IEEE 28th International Symposium on Computer-Based Medical Systems*, 2015, pp. 197–200.
- [30] M. Fan, J. Sun, B. Zhou, and M. Chen, "The smart health initiative in china: The case of wuhan, hubei province," *J. Med. Syst.*, vol. 40, no. 3, pp. 1–17, Mar. 2016. [Online]. Available: <http://dx.doi.org/10.1007/s10916-015-0416-y>
- [31] N. Botts, O. Bouhaddou, J. Bennett, E. Pan, C. Byrne, L. Mercincavage, L. Olinger, E. Hunolt, and T. Cullen, "Data Quality and Interoperability Challenges for eHealth Exchange Participants: Observations from the Department of Veterans Affairs' Virtual Lifetime Electronic Record Health Pilot Phase," *AMIA Annual Symposium Proceedings*, vol. 2014, pp. 307–314, Nov. 2014.
- [32] P. J. del Cid, N. Matthys, C. Huygens, S. Michiels, W. Joosen, and D. Hughes, "Sensor middleware to support diverse data qualities," in *2011 Eighth International Conference on Information Technology: New Generations*, 2011, pp. 673–676.
- [33] S. Sicari, A. Rizzardi, D. Miorandi, C. Cappelletto, and A. Coen-Porisini, "Security policy enforcement for networked smart objects," *Computer Networks*, vol. 108, pp. 133 – 147, 2016.
- [34] J. Singh, T. F. J.-M. Pasquier, J. Bacon, R. Diaconu, J. Powles, and D. Eyers, "Big ideas paper: policy-driven middleware for a legally-compliant internet of things," in *Proceedings 17th ACM/IFIP/Usenix Middleware*, ACM. ACM, December 2016. [Online]. Available: <http://dx.doi.org/10.1145/2988336.2988349>
- [35] I. Caballero, I. Bermejo, L. Parody, M. T. Gómez-López, R. M. Gasca, and M. Piattini, "SL4dq-i8k: Acuerdos a nivel de servicio para calidad de datos en intercambios de datos maestros regulados por iso 8000-1x0," in *X Jornadas de Ciencia e Ingeniería de Servicios (JCIS 2014)*, 2014, pp. 157–166.
- [36] "Towards a service architecture for master data exchange based on {ISO} 8000 with support to process large datasets," *Computer Standards & Interfaces*, pp.–, 2016.